DOCUMENT RESUME

ED 355 807 FL 021 072

AUTHOR Basabas-Ikeguchi, Cecilia
TITLE Analysis of Reading and Listening Comprehension Skills in Different Language Environments.
PUB DATE 88
NOTE 32p.; Master's Thesis, Dokkyo University (Japan).
PUB TYPE Reports - Research/Technical (143) -- Dissertations/Theses - Masters Theses (042)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cloze Procedure; Comparative Analysis; *English (Second Language); Foreign Countries; High Schools; High School Students; *Language Proficiency; Language Tests; *Linguistic Theory; *Listening Comprehension; *Reading Comprehension; Second Language Learning
IDENTIFIERS Unitary Theory

ABSTRACT

The unitary competence theory claims the existence of a general language proficiency that makes the global evaluation of language skills possible and accounts for the high correlations between language tests of whatever form and modality. This study aims to test the theory by investigating the correlation between reading and listening comprehension of students under different linguistic environments. The cloze method was used to measure the reading comprehension of students in a second and foreign language setting, as well as to predict the degree of development of these language skills. The study also investigated whether differences in exposure to English as a Second Language (ESL) and English as a Foreign Language (EFL) yield differences in the degree of development of the language skills, and prove which of the two kinds of exposures leads to greater or lesser skills differentiation. The low correlation between the test rests leads to a question of the unitary competence theory. The results suggest that the cloze system may be a predictor of the differentiated degree of language skills development. Of the two groups taken as samples, the performance of the ESL students in the tests reveal that there is greater tendency for the listening skills to be specialized more than reading comprehension, or vice versa, while exposure to EFL may lead to less differentiation. Contains 37 references. (Author/LB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED355807

TITLE PAGE

ANALYSIS OF READING AND LISTENING COMPREHENSION SKILLS IN DIFFERENT LANGUAGE ENVIRONMENTS

Dr. Cecilia Basabas - Ikeguchi
Faculty of Foreign Languages
Dokkyo University, Japan

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Cecilia Basabas-Ikeguchi
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 021 072

Reading and Listening Comprehension Skills Development in Different Linguistic Environments - in Relation to the Cloze -

Two competing theories have dominated the field of language testing research which seek to demonstrate and explain the structure of language proficiency and measure proficiency levels using external validating criteria. Empirical evidences support the claim of two extreme alternatives: at one end of the theoretical continuum is the claim that the language learner seems to acquire an "overall proficiency" which operates in all the language skills, and this unitary competency leads to the hypothetical possibility of a single unitary factor that accounts for all the reliable variance in all tests of "whatever mode, whatever form". (Oller 1978, 1980, 1983, 1985) The existence of a "general language proficiency" is evidenced by the generally high correlations among a large variety of language competence variables. At the other end of the pendulum is the divisible competence hypothesis that argues that language proficiency can be divided into separate components or several aspects of them. The proponents of the latter claim the possibility of an accountable number of distinct elements of language competencies which prevent any sort of global evaluation. (Lado 1979, 1986) Somewhere in between are the arguements raised by Carroll that there is a "general language ability" , but at the same time, that language skills have some tendency to be developed and

specialized to some degrees, or at different rates, so that different language skills can be separately recognized and measured (1983). Carroll advocates a compromise between the two extremes, preferring a solution that allows a general factor plus various more specialized factors associated with particular skills, components of knowledge and the like.

This study is concerned with the extent to which the unitary competence theory has been explored in ESL and EFL study in relation to the cloze method which has been used to measure the English reading comprehension of none-native English learners. There is wide support from linguistics as well as from language testing research on the use of cloze procedure as a measure of reading comprehension. Its advocates, Oller in particular, claim that this test type is one of the integrative tests (1983) which measure global skills and correlate highly with other integrative tests, especially that of listening comprehension. The most important arguement in its favor is that it requires the FL or SL student to perform a task which is not unlike that which native speakers do in sending and receiving messages.

Recent researches on the use of the cloze test have focussed on the question of construct validity: they have examined the extent to which cloze deletions are capable of measuring comprehension through language abilities beyond the knowledge of the sentence-level grammatical structure. On the one hand there are those who argue that cloze tests are not sensitive to context beyond the sentence. Alderson (1980) systematically varied the size of fixed-

deletion ratios across passages of differing difficulty and across groups, and found no evidence that increase in context affected the subjects' ability to complete the items successfully. Shanahan, Kamil and Tobin (1982) constructed fixed-ratio cloze tests from intact passages, randomly scrambled passages and passages in which the original sentences were extracted and inserted into unrelated texts. They found no significant difference in their subjects' performance across these three conditions. The conclusion reached was that cloze tests do not measure the integration of intersentential information integration, and it seems to be unreasonable to use and interpret cloze in the classroom as a global measure of reading comprehension. Similarly Alderson and Porter (1983) report an empirical finding that cloze tests are either only negligible or else not at all sensitive to constraints ranging beyond 5-10 words of context on either side of the blank. Charles Alderson poses a number of penetrating questions about the validity of cloze tests. He argues that all cloze tests measure the same skill or complex of skills and that in general cloze tests may not be as valid as what it hadbeen previously claimed to be. Alderson's findings claim that "This is not to assert that cloze tests are incapable of testing more than the comprehension of the immediate environment, but that as a test, the cloze is largely confined to the immediate environment of a blank. ..." Moreover the fact that the high correlation of the semantically acceptable scoring method with measures of ESL proficiency, which is designed to be relatively insensitive to long-range contextual constraint, correlated highly with the exact -scoring procedure, appears to add support to the thesis that cloze is essentially clause-bound.

On the other hand, another group of researches have found support for the claim that cloze items are capable of measuring textual relationship beyond the level of sentences. Chihara, Oller, Weaver and Chavez-Oller (1977) using a fixed-ratio deletion procedure with intact and scrambled texts found that both native and non-native speakers of English performed much better on cloze items in intact texts than they did on the same items in the scrambled texts. In addition they found that the percentage of closures was higher for groups with higher language proficiency (native speakers and advanced EFL students) than for groups of lower proficiency (beginning and intermediate EFL students). Bachman (1982) using a rational deletion procedure with an intact passage, found that deletion types requiring context beyond the clause for closure loaded on different factors rather than did deletions which depended only upon the context of the clause in which they occurred. The arguements raised in the study are based on the assumption that within a given text more words function syntactically than cohesively, a random deletion procedure would tend to sample a large portion of clause-bound words, and therefore appear to be measuring only lower-level skills. But the findings that the high-level proficiency non-native speakers performed very similarly to native English speakers on both syntactic and cohesive cloze items, while low proficiency English sepakers performed significantly poorer on cohesive items, suggest that cloze passages using a rational deletion procedure can be used to measure textual relationship beyond clause boundaries. More recently the experiment by Chavez-Oller, Chihara and Weaver and Oller (1985), using

previous data (1977) indicates that the performance of the subjects (both native and non-native English speakers) differed mostly between the scrambled and syntactic conditions. A linguistic analysis demonstrated that virtually all the items depended on intersentential context for successful closure.

The performance of EL1 and EL2 respondents on Bachman's innovative experiment with several deletion types (1985) further throws light on the question of whether cloze measures higher order skills. Bachman, in this study, was primarily concerned with constructing valied and reliable cloze tests for reading development; however, the different responses of the EL1 and EL2 subjects on the different deletion types support the claim that cloze items are sensitive to discourse across sentences. These variations in performance of native speakers as well as those of the non-native speakers might indicate that cloze tests do measure higher-order skills.

Along with the external component of construct validation, a substantial amount of research on cloze items has found evidence for some cloze items as measures of textual competence (Shanahan and Kamil 1982) and others as measures of grammatical competence. (Weaver and Oller 1985, Lado 1986 and Markham 1985). Much of the research on cloze procedure mentioned above consisted of the analyses of cloze and other language tests which were all shown to correlate to some degree (Shikkani 1986) or to load in one general factor (Oller, 1983). A single factor which accounts for most of the variance in language tests has been reported in many studies in which cloze

has been used. (Oller and Hinofitis, 1980) and Alderson (1980, 1983) Although Alderson does not refer to a general factor, the assumption that cloze measures a rather homogenous group of lower-order core proficiency skills seems to underlie much of his research.

The present experiment was conducted to partially determine the discriminative power of the cloze test (scored by the exact and the acceptable scoring methods) and its ability to predict other language skills, specifically that of listening comprehension. The basic concern of this study is to determine whether differences in the degree of exposure to EL2 - as a foreign language and as a second language - would result in differences in the degree of development between the language skills, particularly those of listening and reading comprehension.

Experimental Method

Subjects

A total of 230 subjects, consisting of two groups of high school students, were taken for the experiment. The first (EFL) group consisted of 116 first year senior high school students in Japan; the second (ESL) group consisted of 114 third year high school students in Manila. The age range of the population was from 15-16. Since the students in group 1 were enrolled in a high school course where no attempt had been made at ability grouping, it is assumed that they were representative of the larger population from which the samples were drawn. Meanwhile the students in the second group were selected

from one high-ability class, one average-ability class, and one low-ability class, as per the structure of the school they are enrolled in.

It is at this point where a careful word of distinction is necessary to explain what may be a verbal quibble concerning the linguistic environments of the subjects of this paper: English as a Second Language and English as a Foreign Language. By English as a Foreign Language for the (hereon referred to as the) first group of Ss, is meant that English is taught as a school subject or on an adult level solely for the purpose of giving the students a foreign language competence which he may use in one of several ways - to read literature, to read technical work, to listen to the radio, to understand a dialogue in the movies, to use the language for communication possibly with transient English or Americans. In general the students' motivation is for the practical use of English, while some of them realize at the same time that the language is an effective key to the understanding of another culture.

The term English as a Second Language is used to describe the situation of the (hereon referred to as) second group of Ss where English becomes a language of instruction in schools. It is often fundamental to the life and functioning of the country rather than a region where it has a foreign language status. But another word of caution has to be pointed out here: it is that in situations involving both English as a Second or Foreign Language, the word second appears to be an understatement. In the case of the language environment of the second group of Ss, the school child has had contact with English as

early as right after birth, but the first language comes first - or simultaneously - with another regional dialect. This means that English learning is complicated by interference from two sets of language structures instead of just one. So even between English learners in an L2 setting, there are differences between learning motivations, type and amount of exposure to the target language, as well as points of emphasis on the L2 learning at home and in school, as much as there are critical differences between first and second language learning as discussed by Elley and Mangubhai (1983).

Experimental Materials and Design

The cloze passage used in the research was adopted from the text High Road to English, a high school reader. The selected passage contained approximately 350 words, the title of which is "Four Boys and the French Cave of Lascaux". The close itself was constructed by deleting every seventh word in the passage for a total of 50 blanks. Two sentences were left unmutilated at the beginning of the passage and one at the end to provide a complete context.

The passage had a Flesch readability (1948) score of 88.9, ranking it as "easy" and a human interest score of 50, indicating that the article is "highly interesting". The Dale-Chall formula (1948) ranked it as appropriate for the fourth or fifth grade, with a score of 4.8. This indicates the reading grade of the student who could answer one-half of the test questions correctly.

Both the SMOG and the Fry formulas gave approximately the same readability grade for the passage.

For the listening comprehension test, a TOFEL exam used in a Japanese school was used with the Ss. Under the present conditions, it was assumed that a TOFEL test would be a more valid measure for both the Ss rather than a TESOL exam which would be more difficult for the first group of samples.The listening comprehension test totalling 65 points consisted of seven parts. In the first part, the Ss were asked to select one logical answer from among three choices given. In the second part, the students heard ten short questions that follow a recorded conversation. Then they were asked to select one logical answer from among the three choices given. Part Three was understanding of three recorded passages which were played twice. The students were instructed to answer questions relevant to the passages by choosing from choices of three items each. In Part four, the Ss were asked to decide if the meaning expressed by each statement is logical or not. Part five consisted of a much longer passage than those in Part three. The students heard the passage twice, after which they heard five short sentences. They were asked to decide whether the sentences they heard were relevant to the passage or not. Part six consisted of ten recorded short questions which the Ss heard only once. They were then instructed to select one appropriate answer from three given choices. Part seven consisted of question-answer form for daily conversation. The students were asked to select from three choices the answer that best fits the given situation.

The subjects were tested in groups of 40 persons or more according to

the class size. Each of the subjects took the text of the close passage on the first day and the listening comprehension test on the second day. The tests were administered in a regular class time under regular classroom conditions.

For the first group of students, instructions for the cloze test were read aloud in English and then explained in Japanese by the Japanese examiner. For the second group, instructions were first read aloud by the researcher and were read silently by the students themselves. They were instructed to read over the passage quickly for general meaning, then to fill in the blanks, and finally to check their answers. The LCT was recorded on tape by a native speaker, and the students were allowed to listen to each item of the test, and answer the questions within the time limit indicated in the tape. The subjects were given 30 minutes to answer the cloze test and a fourty-minute class period to answer the listening comprehension test.

Results and Discussion

Scoring

Researches with native English learners have explored the various methods of scoring cloze tests and have shown that the most economical and reliable scoring method is the exact-word scoring method. Of the prominent researches are those conducted by Taylor, Runkin, and Ruddell and Bormuth (1976), and Brown (1983). Moreover, the use of the exact-word scoring method yields to a slightly different result with native speakers than the contextually acceptable scoring method, which is more tedious and requires

more effort. (Oller and Conrad, 1977, 1983, 1985)

Researches with non-native English learners however, such as those of Oller (1977), and Alderson (1983) show that the correlations between cloze and external validating criteria are significantly superior when the contextually acceptable system of scoring is used. The most important arguement on the use of the semantically acceptable scoring system was raised by Oller in defense of cloze as a measure of intersentential comprehension. His findings indicate that scoring for contextual appropriateness would probably have augmented the difference across normal and scrambled conditions. "Further since a contextually appropriate scoring could only be achieved by using the original textual order as the criterial basis, it would have tended probably, to create a priori bias against the scrambled condidtions." The exact-word scoring method is too strict an application for the non-native English learners. Furthermore the task of guessing the exact word placement is not necessarily a language skill (Oller, 1983). Other researches who have experimented with the cloze system on FL2 have often preferred scoring systems that give credit for contextually acceptable answers.

Since the concern in the present study is to test the result of using these methods for ESL Ss and EFL Ss, both the exact-word and the acceptable scoring methods were used. The first method, in spite of its handicap for use with non-native samples, was nevertheless carried out to provide a basis of comparison as to which of the two systems of scoring will yield to higher

correlations with the TOFEL listening test used. Assuming the arguements in its favor- notwithstanding the difficulty of application - that the contextually acceptable scoring method yields higher correlations and is rationally acceptable with EL2 subjects, the second scoring method was applied with a native speaker's judgement.

The Ss were instructed to use only one word to fill in the blanks. In spite of this, however, there were instances when the respondents used more than a one-word response. In the first scoring, these multiple responses were counted incorrect. In the second scoring, however, these answers were judged correct or incorrect on the basis of their acceptability in the context given in the passage. Because of this and the re-scoring of the multiple-responses, the second scoring method is said to have been less strict than the first method.

Test Analysis

To give an indication of the discriminative power of the cloze test, the data were analyzed by determining the means, standard deviation, mode and the range. Table 1 summarizes the results of the first and second scoring of the cloze test plus the reliability coefficient scores.

Table 1: *Cloze Test Results with Reliability Coefficients*

	FIRST SCORING		SECOND SCORING	
	EFL	ESL	EFL	ESL
No. of Data	116	114	116	114
Mean	19	28	25.16	40.45
Mode	18	30	29	39
Std. Dev.	19	15	8.76	5.32
Max. Score	36	40	43	49
Min. Score	5	20	6	28
Range	31	20	37	21
Reliability Estimates:				
		Cron-Al	.907	.801
		K-R 20	.907	.801
		K-R 21	.854	.742

CLOZE TEST RESULTS

The difference in the means obtained by the two groups after using the second scoring system is much greater than that obtained after using the first scoring method. The scores of the first group of Ss for the second scoring, are on the whole, only slightly higher than their scores after the first scoring had been done. The mean scores obtained by this group (the EFL Ss) in both scoring systems have been influenced by the very low base scores which are 5 and 6. As far as the mode is concerned, however, majority of the students in this group who scored below 50% in the first scoring - notice that

the mode is 18 - performed above 50% in the second scoring - with the mode of 29. The table likewise indicates the wide spread of some of the scores of the EFL samples - with ranges of 31 and 37 - compared to the other group - with a score range of 20 - 21.

The scores of the second group of Ss (ESL) are, on the whole, much higher in the second scoring than in the first. Most of the students who already scored above 50% in the first scoring obtained even higher when the semantically acceptable scoring method was applied. Reliability estimates used for the cloze test were the Cronbach Aplha, Kuder-Richardson Formula 20 and Kuder-Richardson Formula 21. All the estimates gave .8 and .9 coefficients for the Japanese Ss and slightly lower coefficients for the Filipino students.

LISTENING COMPREHENSION TEST RESULTS

Table 2 reports the means and standard deviation together with the reliability estimates of the listening comprehension test used. The results indicate that in contrast to the cloze test results, the spread of the scores of EFL Ss is much narrower than that of ESL Ss. On the other hand, the raw scores of the latter is much higher than those of the former, notice the 55.7 and 45.5 mean scores respectively, with a lower STD in the ESL group (5.5) than in the other (8.1)

Table 2: *Listening Comprehension Test Results with Reliability Estimates*

	EFL	ESL
No. of Data	116	114
Mean	45.53	55.70
Mode	49	57
Std. Dev.	8.117	5.526
Max. Score	65	63
Min. Score	27	19
Range	38	42
Reliability :		
Cron-Al	.832	.805
K-R 20	.832	.805
K-R 21	.808	.746

For both tests, the performance of the ESL students was higher than that of the EFL samples. This is indicated by the higher mean performance as well as the smaller standard deviation of the former.

Significantly, one student from the EFL group got a perfect score while the highest score in the second group was two points lower. Meanwhile, the single score which best represents the performance of Group 1 is lower than that of Group 2, with the performance modes of 49 and 57, with the mean scores being 45 and 55, respectively. The reliability estimates were obtained by using the Cronbach Alpha, KR 20 and KR 21 formulas. The coefficients obtained are in .8 range for the first group of subjects; while for the other group, they were found to be within the .7 and .8 range.

On the whole, the test results obtained from the two groups of samples indicate a low correlation, although positive. The slightly higher correlational results of the ESL Ss indicate that the cloze may not or does not necessarily predict other language performance skills.

Fig. 1 - 4 present the results of the correlational analysis done on the cloze and listening comprehension tests.

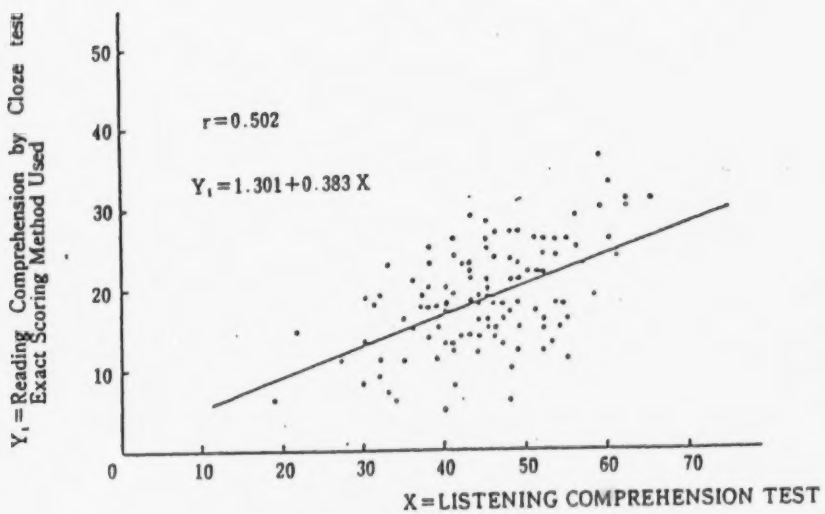

Fig. 1: EFL =LCT and RCT (Exact Word Scoring Method) S_1

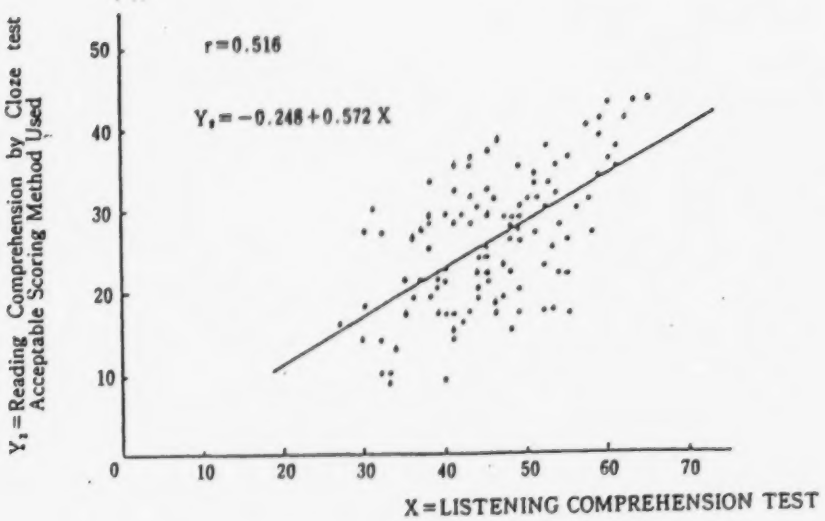

Fig. 2: EFL = LCT and RCT (Acceptable Scoring Method) S_2

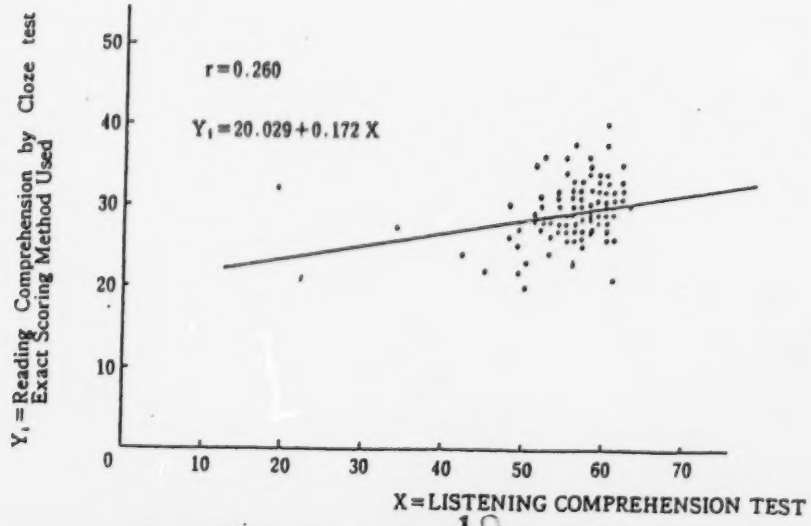

Fig. 3: ESL = LCT and RCT (Exact Word Scoring Method) S_1

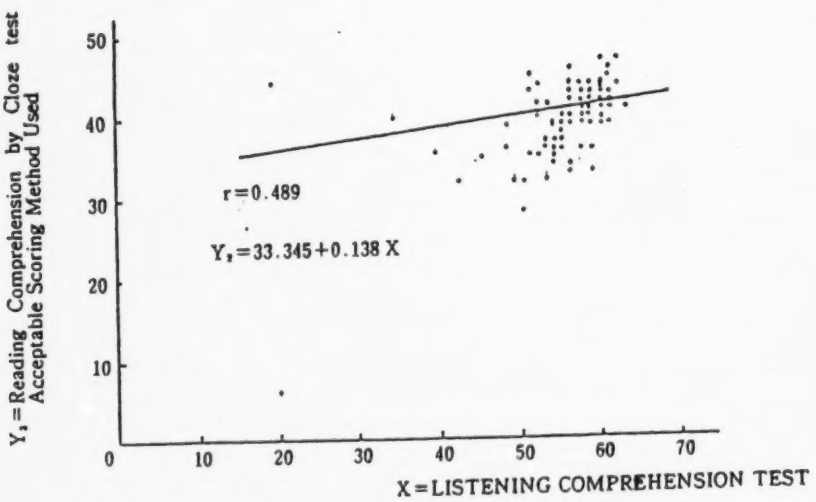

Fig. 4: ESL = LCT and RCT (Acceptable Scoring Method) S_2

Discussions of Results

Contrary to expectations, the results of the study yielded very low, although positive correlations between the two tests. Comparatively, the results of the tests of the EFL Ss gave a higher correlation than those of ESL Ss. Furthermore, the acceptable scoring method (S_2) in Fig. 2 and Fig. 4 yielded to a slightly higher correlation with the TOFEL test performance of both groups than did the exact scoring method (S_1) in Fig. 1 and Fig. 3. Finally, the inter-group correlational difference with the application of the two scoring methods varies and that such a difference is much higher in the second group

(Fig. 3 and 4) than the difference in the scores of the first group (Fig. 1 and 2). With the application of the semantically acceptable scoring method, a higher correlation with an external criterion - such as the TOFEL - was obtained. It should be noted that the results were obtained by using the PPM Correlation formula that makes use of data from raw scores.

Statistically, the data reported here are interesting; first in that it is likely that the high schools students who were taken as samples comprised a relatively homogenous group of individuals, and this curtailment of range, specifically of age, may have resulted in spuriously low correlations.

Another factor that might have greatly influenced the very low correlation of the test results in the second group is the presence of "outliers". Fig. 3 and 4 show the existence of the two greatly deviant scores which fall below the lowest score in the listening comprehension test. The reason for the existence of these "outliers" can best be studied through a careful case study, which could not be included in this present study due to apparent reasons of space and time. The clustered test scores may have resulted in the low correlation.

As regards cloze test scoring methods, a tentative conclusion that can be drawn from the data of this research is that between the exact and contextually acceptable scoring systems, the latter yields higher correlations with other tests than the former. This confirms the results of some previous researches,

especially those of Darnell, et al (1976) and Vollmer and Sang (1983) that the contextually acceptable scoring method is superior in terms of correlational results when applied to samples of non-native speakers.

The higher mean scores and higher correlation results with the listening comprehension results of both groups of samples when the semantically-acceptable scoring system was used support the claim that it is the rationally justifiable with subjects to whom English is a second or a foreign language.

The difference in the performance of the two groups of non-native English learners present striking insights not only into the relationship between listening comprehension and reading but also tend to differentiate students' exposure to English as a second or foreign language. A number of factors may be thought of, the most influential of which may be the degree to which they are exposed to and use English in the school, at home and in the environment.

Although it is not the aim of this paper to explore the complex process of reading in an SL or FL setting, it is interesting to look into how the skills of reading and listening to English as FL or SL is organized and how they develop or are developed. But in answering this question, one is really asking, as Carroll suggests, whether these skills are attained at the same rate, or attained separately, at different rates (1983). It is possible to trace the actual rate of skill development by subjecting the samples to certain observational

criterion over a period of time and using the refined factor analysis. A high correlation between a pair of skill variables would indicate that to the extent the learner has attained one of these skills, he is most likely to have attained the other skill at about the same relative extent. The findings of this research support this claim. The comparatively high correlation between the cloze and listening test results of the Japanese students indicates that exposure to English as a foreign language has made possible the fact that the language skills have been tapped and used at an equally same degree at this stage.

This implies that the cognitive skills learned in decoding the messages in both the audio and written tests have advanced to a certain extent that such advancement in the listening skills is accompanied by advancement in the other skill, or vice-versa (Carroll, 1983). But what does this indicate in relation to the samples to whom English is a second language?

For the second group of population, with English as a medium of instruction in school, and for most of whom English is the major language at home and in the environment, the low correlational test results may indicate a fairly unequal development of the receptive skills of audio-visual comprehension such that the skills in listening comprehension have advanced to a certain degree while the other skills of reading comprehension have not. It may happen that in a second language environment, the students are exposed more to aural English, the language they first come in contact with, and more because it is a comfnon language in mass media, and even more when it is

used as a medium of instruction. Or it may be the reverse; it is hard to determine at this point which skill has been developed more and which has advanced less. The close test results can not point out to this. The findings simply indicate that there is a gap in the level of skills advancement.

But just how exactly these language skills develop in order is unknown in SL learning (Verhoeven, 1990) and Nagle and Sanders (1986). The extensive discussions of James and Sticht (1982) on the relationship between listening and reading provide helpful insights on this area of research. Although the article discusses mainly issues on reading and listening in relation to L1, it nonetheless throws light on this study, assuming a Unitary Competence and A General Language Proficiency Theory. Besides, as Oller (1977) has earlier pointed out in relation to the cloze, reading and listening tasks require the FL and SL learners to perform a task which is not unlikely to what native speakers do in receiving and sending information. Furthermore, studies on ESL2 learners suggest that L2 acquisition in naturalistic context has much in common with first language acquisition. (Elley and Mangubhai, 1983).

The most important arguement raised in the article of James and Sticht is that the development of the skill of auding is a pre-requisite to the development of the skill in reading. The reading potential concept developed by James and Sticht indicates that auding and reading performance will be more highly correlated after some skills of learning to read has developed. This may have been the case with the EFL students taken as samples. However, this does not mean that the ESL learners have not learned to read better than the EFL - quite

the contrary, as seen in higher reading test results. The higher mean scores and lower STD in the listening test- compared to the other group and to their own (ESL) reading scores is not unexpected considering the fact that English is a near-native language; it is "common", interchangably used with - or even more than - the native tongue. This being the case, they would have had the potential for higher auding skill. And according to James and Sticht's claim on the transfer effect, pupils highly skilled in auding would have (greater) potential to become highoy skilled in reading, which is proven by higher reading scores.

But the correlationally lower test results of the ESL samples support another issue raised by the article that although exposure to spoken language is necessary for reading to take place, auding-to-reading transfer may not necessarily occur in the auding, then reading order. It may have been the reverse. Congruent with this arguement, the results with the EFL Ss indicate that a reading first, then, auding transfer may have been possible because the age of the Ss (3rd year JHS) mean that they may have had the time to learn the peculiarities of written English more -which reflects strongly the emphasis on grammar and translation method very common in Japanese schools.

On the whole, the relatively low correlation in the test results of both the EFL and ESL students lead to a question on the unitary competence hypothesis postulated by Oller. "If the components of language competence

function more or less similarly in any language-bared task, high correlations would be expected between language test scores of all sorts, whatever modality, whatever format."(Oller, 1983) Given the limited conditions and the limited number of samples for this study, what can be tentatively stated is that the different language skills or listening and reading are learned together and developed in different degrees according to whether exposure to English is as a FL or as an ESL, and yet there is a general tendency for some language skills to be developed and specialized to some degree than the others and can be measured and recognized as distinct from other language skills. Thus the results of a cloze test may be, but not the only predictor of the listening comprehension ability of English as the foreign or second language learners.

As regards cloze test scoring methods, another tentative conclusion that can be drawn from the data of this research is that between the exact and contextually acceptable scoring systems, the latter yields higher correlations with other tests than the former. This confirms the results of some previous researches (Alderson, 1983 and Bachman 1982, 1985) that the contextually acceptable scoring method is superior in terms of correlational results - especially with L2 samples - compared to other scoring systems.

Because of the very limited time involved in this study, the relatively small number of population and the limitations of only one cloze test and one listening comprehension test used, the findings and conclusions arrived at are only inclusive of the present data. Other important areas of concern have not

been addressed in this study, and are open to further research.

First, a response frequency analysis need be carried out for each item of the cloze test and / or listening comprehension test as completed by each group of subjects. The main question to be addressed here is whether or not the responses of the EFL students would be significantly different in kind as well as in number from those of the ESLstudents.

Second, a further experimentation with another improved cloze test of a different difficculty level need to be conducted and administered to a bigger, more heterogenous population than those used in this experiment, and that the test be further correlated with other external validating criteria. Why the second group of Ss scored lower in spite of the use of the TOFEL exam in listening need further investigation.

Third and closely related to the second, the low reading test results point out to a need to look into the reading problems that go with EL2 study. The complex mechanism involved in the acquisition of reading in EL2 discussed in Verhoeven (1990) runs along parallel lines with the discussions of James and Sticht (1982) that emphasize that children should be helped to build up oral skills in EL2 first before reading instruction is started. In the case of the samples used in the present study, instructional reading materials are presented simultaneously with audio instruction, although environmental exposure to auding starts earlier in the ESL Ss than in the EFL Ss.

Then too, if the low correlation between the tests indicate a gap in the development of the language skills of the non-native English students, a more systematic search be started for factors related to increases in English language proficiency at various stages of ESL and EFL study under diverse circumstances of language learning as those of Japan and the Philippines. (O'Mally and Chamot, 1985) and Vollmer and Sang (1983). This experiment has identified a gap in the development of the language skills and leaves room for the next task which is for the researches and for those in the field to reduce the obvious disparities in competence levels as a result of such gap.

REFERENCES

Alderson, J.C. (1980). Native and Non-native Speaker Performance on cloze test, *Language Learning*, 30, 59-76.

Alderson, J. C. (1983). The Cloze Proficiency and Proficiency in English as a Foreign Language. In John Oller (Ed), *Issues in Language Testing Research* (pp. 205- 217). Rowley, MA: Newbury House.

Anderson, J. (1977). A Technique for Measuring Reading Comprehension , *Language Learning* , 21, 178 - 181.

Bachman, L. (1982). The Trait Structure of Cloze Test Scores, *TESOL Quarterly*, 16, 61-67.

Bachman, L. (1985). Performance on Cloze Tests with Fixed-Ratio and Rational Deletions, *TESOL Quarterly*, 19, 535 - 551.

Belasco, S. (1965) Nucleation and Audio-lingual Approach, *Modern Language Journal*., 48, 217 - 223.

Belasco, S. (1978). Feasibility of Learning English in a SL Unicultural Situation, In Pimsleur *Psychology of Second Language Learning*. (1 - 10). Cambridge: University Press.

Brown, J. (1983). A Closer Look at Cloze. In Oller, J. (Ed.), *Issues in Language Testing Research* (pp. 237 - 250). Rowley, MA: Newbury House.

Carroll, J.B. (1977). Development Parameters of Reading Comprehension. In J.T. Guthrie (Ed), *Cognition, curriculum and comprehension* (pp

11-15). Newark, DE:International Reading Association.

Carroll, J.B. (1983). Psychometric Theory. in John Oller (Ed), *Issues in Language Testing Research*: (pp. 80 - 107): Newbury House.

Chapelle, A. and Abraham, R. (1990). Cloze Method: what difference does it make? *Language Testing*, 7, 121- 146.

Chastain, K. (1976). *Developing Second Language Skills*. Boston: Houghton Miffin.

Chavez-Oller, Chihara, T., & Oller, J. (1985). When are Cloze Items Sensitive to Constraints Across Sentences? *Language Learning*, 35, 181 - 202.

Chihara, T., Oller, J.W., K-Weaver, & Chavez-Oller, (1977). Are Cloze Items Sensitive to Constraints Across Sentences? *Language Learning*, 27, 63 - 73.

Dale, E. & Chall, J.S A Formula for Predicting Readability. *Educational Research Bulletin*, 1948, 27, 11-20, 37 - 54.

Durrell, D.D. (1969). Listening Comprehension VS Reading Comprehension. *Journal of Reading*, 12, 455-460.

Elley, W. & Mangubhai, F. (1983). The Impact of Reading on Second Language Learning. *Reading Research Quarterly*, 19, 53 - 67.

Koike, I. Ed. (1978). *The Teaching of English in Japan*. Tokyo: Eichosa Publications.

Lado, R. (1980). *Language Testing*. (pp. 315 - 333). London: Longman Press.

Lado, R. (1986). Analysis of Native Speaker Performance in Cloze Tests.

Language Testing, 3, 130 - 153.

Markham, P. (1985). The Rational Deletion Cloze and Global Comprehension in German. *Language Learning*, 35, 423 - 430.

Nagle, S. & Sanders, S. Comprehension Theory and Second Language Pedagogy. *TESOL Quarterly*, 20, 9 - 24.

Oller, J. Ed. (1983). *Issues in Language Testing Research.* (205 - 217). Rowley: MA: Newbury House.

Oller, J.W. & Hinofitis, F.D. (1980). Two Mutually Exclusive Hypotheses about Second Language Ability. In Oller and Perkins (Eds.), *Research in Language Testing*, Rowley:MA: Newbury House.

Oller, J. (1979). Pscychology of Language and Contrastive Linguistics. *Foreign Language Journal*, 12, 299-309.

Oller, J. (1972). Assessing Competence in ESL: Reading. *TESOL Quarterly*, 6, 321 - 323.

O'Mally, Chamot, A. et al. (1985) Learning Strategies Used by Beginning and Intermediate ESL Students. *Language Learning*, 35, 21-46.

Perkins, K. (1983). Semantic Constructivity in ESL Reading Comprehension. *TESOL Quarterly*, 17, 19 - 27.

Shanahan, T., Kamil, M. & Tobin, A.W. (1982) Cloze as Measure of Intersentential Comprehension. *Reading Research Quarterly*, 17, 229-255.

Shikhani, M. & Hanania, E. (1986). Interrelationships Among Three Tests of Language Proficiency. *TESOL Quarterly*, 20, 97 - 122.

Sticht, T. and James, J. (1984). Listening and Reading. In P. Pearson, R.

Barr, M. Kamil &Mosenthal, P. (Eds.), *Handbook of Reading Research.* New York: Longman. (293-318).

Towsend D.J., Carrithers, C., & Bever, T.G. (1987) Listening and reading process in college-and middle school-age readers. In R. Harowitz & S.J. Samuels (Eds.), *Comprehending Oral and Written Language* (217 - 242). New York: Academic Press.

Verhoeven, T. (1990). Acquisition of L2 reading, *Reading Research Quarterly*, 25, 91 - 112.

Vollmer, H. & Sang, F. (1983) Competing Hypotheses about Second Language Ability. In John Oller (Ed.), *Issues in Language Testing Research* , (pp. 29 - 79). Rowley MA: Newbury House.

Winitz, H. Ed. (1981) *Native Language and Foreign Language Acquisition.* New York: Academy of Science Press.